Wax

i. 632

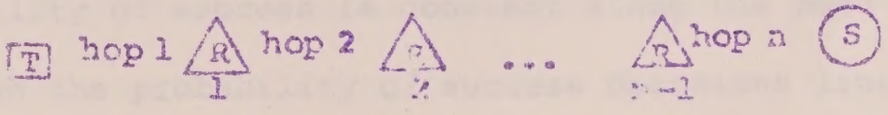
ARPA Network Information Center
Stanford Research Institute
Menlo Park, California 94025
NIC#13632
Packet Radio Temporary Note 7

Network Analysis Corp.
12 January 1973

COMPARISON OF HOP-BY-HOP AND END-TO-END ACKNOWLEDGMENT SCHEMES

I. Introduction and Summary

We consider the case where (n-1) Repeaters separate the packet radio Terminal from the Station. Assuming that the terminal is at a distance of "one hop" from the first repeater one obtains the following n-hop system:



A simple model is used for evaluating the total average delay that a packet encounters in the n-hop system when using hop-by-hop and end-to-end acknowledgment schemes. When the end-to-end acknowledgment scheme is used, every repeater transmits the packet once. If the packet does not reach the station then retransmission starts from the terminal. The acknowledgment is sent from the station. In the hop-by-hop scheme repeaters store and retransmit the packet until positively acknowledged from the next repeater stage. Thus, one obtains "hop-by-hop" transmission.

The operation is so that a terminal, or a repeater in the "hop-by-hop" case, transmits the packet and if an acknowledgment does not arrive within a specified period of time, it

retransmits the packet. The waiting period is composed of the time for the acknowledgment to arrive when no conflicts occur plus a random time for avoiding repeated conflicts.

Two different schemes for end-to-end acknowledgment and one scheme for hop-by-hop acknowledgment are studied. Curves for the total average delay as a function of the number of hops and the probability of successful transmission per hop are obtained. Two cases were considered: one in which the probability of success is constant along the path and another in which the probability of success decreases linearly as the packet approaches the station. Finally, channel utilizations are compared when using a slotted ALOHA random access mode of operation.

It is demonstrated that the hop-by-hop scheme is superior in terms of delay or channel utilization. This conclusion becomes significant when the number of hops increases or when the probability of successful transmission is low. For example, in a five hop system, if the probability of success per hop is 0.7 then the total average delay is 12.5 and 53 packet transmission times for the hop-by-hop and end-to-end acknowledgment schemes, respectively. The functional capabilities of the hardware required for using the schemes considered can be found in Packet Radio Temporary Note #5.

The model used is based on ASS Note 9 by L. G. Roberts and ASS Note 12 by L. Kleinrock and S. S. Lam. The model is simplified, however, by assuming that the probability that a packet is blocked is the same when the packet is new or has been blocked any number of times before. Although the more general equations could have been written, the numerical solution is rather elaborate (see ASS Note 12) and seems unnecessary for this comparative study. It is further assumed that the probability of being blocked in different hops are mutually independent. By total delay is meant from the time that the first bit is transmitted by the terminal until the time that the last bit of the packet is correctly received by the station.

II. Delay Equations

Notation

P = propagation delay per hop in one direction

T_f = transmission time of information packet

T_b = transmission time of acknowledgment packet

σ = rate of packets offered to receiver

γ = rate of packets with retransmission offered to receiver

q_i = P[successful transmission in hop i]

A = successful transmission delay

Digitized by the Internet Archive
in 2025 with funding from
Amateur Radio Digital Communications, Grant 151

https://archive.org/details/temporary-note-007

R = average waiting time before retransmission is made

$\bar{X}$ = average waiting time beyond the minimum to avoid repeated conflicts (we assume same for the different schemes considered)

D = E[total delay]

ρ = channel utilization

C = channel capacity

h, s = superscripts for denoting hop-by-hop and end-to-end transmission schemes respectively.

Above quantities with subscript i relate to i-th hop, unsubscript quantities refer to end-to-end or are the same for each hop, depending on the context.

Hop-by-hop acknowledgment

For hop i one has:

$$A_i = P + T_f$$

$$R_i = 2P + T_f + T_b + \bar{X}$$

That is, the repeater waits the transmission time and the propagation time to the next repeater plus the transmission time of the acknowledgment packet and its propagation backward plus, finally, a random time for avoiding repeated conflicts.

R = average waiting time before retransmission is made

$\overline{X}$ = average waiting time beyond the minimum to avoid repeated conflicts (we assume same for the different schemes considered)

D = E[total delay]

ρ = channel utilization

C = channel capacity

h,e = superscripts for denoting hop-by-hop and end-to-end transmission schemes respectively.

Above quantities with subscript i relate to i-th hop, unsubscript quantities refer to end-to-end or are the same for each hop, depending on the context.

Hop-by-hop acknowledgment

For hop i one has:

$$A^i = P + T_c^i$$

$$R^i = 2P + T_c^i + T_p^i + \overline{X}$$

That is, the repeater waits the transmission time and the propagation time to the next repeater plus the transmission time of the acknowledgment packet and its propagation back-ward plus, finally, a random time for avoiding repeated conflicts.

$$E[\text{no. of retransmissions in hop } i] = \sum_{i=1}^{\infty} i(1-q_i)^i q_i = \frac{1-q_i}{q_i}$$

$$D_i = A_i + \frac{1-q_i}{q_i} R_i$$

$$D^h = \sum D_i = n(P + T_f) + (2P + T_f + T_b + \bar{X}) \sum_{i=1}^{n} \frac{1-q_i}{q_i} \qquad (1)$$

End-to-end acknowledgment

$$A = n(P + T_f)$$

$$R = n(2P + T_f + T_b) + \bar{X}$$

$$P[\text{successful transmission}] = Q = \prod_{i=1}^{n} q_i$$

$$E[\text{no. of retransmissions}] = \sum_{i=1}^{\infty} i(1-Q)^i Q = \frac{1-Q}{Q}$$

$$D^{sl} = A + \frac{1-Q}{Q} R = n(P + T_f) + [n(2P + T_f + T_b) + \bar{X}] \frac{1-Q}{Q} \qquad (2)$$

We also consider another method of end-to-end acknowledgment in which the terminal retransmits after shorter periods of time due to anticipating a low probability of successful transmission Q. In particular, we examine the case in which the retransmission delay is the same as in the hop-by-hop method. Another motivation for introducing this method is that terminals which are adjusted to operate in an urban area (one hop system) and operate from a rural area will work that way. The total average delay

$$E[\text{no. of retransmissions in hop } i] = \sum_{i=1}^{\infty} i(1-q_i)^i q_i = \frac{1-q_i}{q_i}$$

$$D_i' = A_i' + \frac{1-q_i}{q_i} R_i'$$

$$D_p = \sum D_i' = n(p + T_E') + (2p + T_E' + T_D' + \bar{X}) \sum_{i=1}^{n} \frac{1-q_i}{q_i} \qquad (1)$$

End-to-end acknowledgment

$$A = n(p + T_E')$$

$$R = n(2p + T_E' + T_D') + \bar{X}$$

$$P[\text{successful transmission}] = Q = \prod_{i=1}^{n} q_i$$

$$E[\text{no. of retransmissions}] = \sum_{i=1}^{\infty} i(1-Q)^i Q = \frac{1-Q}{Q}$$

$$D_{ej} = A + \frac{1-Q}{Q} R = n(p + T_E') + [n(2p + T_E' + T_D') + \bar{X}]\frac{1-Q}{Q} \qquad (2)$$

We also consider another method of end-to-end acknowledgment in which the terminal retransmits after shorter periods of time due to anticipating a low probability of successful transmission Q. In particular, we examine the case in which the retransmission delay is the same as in the hop-by-hop method. Another motivation for introducing this method is that terminals which are adjusted to operate in an urban area (one hop system) and operate from a rural area will work that way. The total average delay

for this method is given by:

$$D^{s2} = n(P + T_f) + (2P + T_f + T_b + \bar{X}) \frac{1-Q}{Q} \qquad (3)$$

To write the average delay normalized to the packet transmission time T_f, we introduce the following parameters:

$$\alpha = \frac{T_b}{T_f} \qquad (4a)$$

$$\beta = \frac{P}{T_f} \qquad (4b)$$

$$\gamma = \frac{\bar{X}}{T_f} \qquad (4c)$$

ρ is a function of the channel capacity C, the packet size in bits L, and the distance in miles per hop d; given by

$$\rho = \frac{1}{c} \cdot \frac{d \cdot C}{L} \qquad (5)$$

where c is the propagation speed 168,272 miles/sec in free space. γ is the ratio of the average "extra" time to T_f for randomizing

retransmissions. For example, if a slotted ALOHA transmission system is used with the random "extra" time uniformly distributed over k slots, as in ASS Notes 9 and 12, then $\delta = (k-1)/2$.

The normalized delays are now written as:

$$D^{h'} = (1+\beta)n + (1+2\beta+\alpha+\delta)\sum_{i=1}^{n}\frac{1-q_i}{q_i} \qquad (6)$$

$$D^{s1'} = (1+\beta)n + [(1+2\beta+\alpha)n+\delta]\frac{1-Q}{Q} \qquad (7)$$

$$D^{s2'} = (1+\beta)n + (1+2\beta+\alpha+\delta)\frac{1-Q}{Q} \qquad (8)$$

III. Constant q Along the Path

In this section we present the delays for the case in which the probability of successful transmission is the same for every hop along the path. The curves shown in the figures are for the parameters: $\alpha = 0.5$, $\beta = 0.02$, and $\delta = 2.0$. For example, $\beta = 0.02$ occurs when $d \cong 15$ miles for which the propagation time is 80 μ sec, and when the packet transmission time is 4 msec (eg. 400 bits @ 100 Kb/s).

Note that the curves show the delay as a function of q rather than channel utilization. Thus, they can be used for a slotted or non-slotted random access ALOHA Systems or possibly

other modes of operation. When ALOHA system is used, one can use Figure 1 which shows the effective channel utilization and the channel use (i.e. with retransmissions) as a function of q. For the slotted case they are obviously twice the values for the non-slotted case. The formulae used for Figure 1 for the non-slotted case , $q=\exp(-2\gamma T_f)$, are (similarly for the slotted case):

$$\text{channel use} = \gamma T_f = 1/2 \ln \frac{1}{q} .$$

$$\text{channel utilization} = \lambda T_f = \gamma T_f q$$

Figure 2 shows the normalized delay as a function of q, with n as a parameter. One can see that the delay for the end-to-end acknowledgment schemes grows much faster than in the hop-by-hop scheme. For example, in a 5-hop system, when T_f = 4 msec and q = 0.6 then the average delays are 68 msec, 188 msec, and 472 msec, for D^h, D^{s2}, and D^{s1}, respectively. Alternatively, for a 5-hop system and T_f=4 msec, assume that an acceptable average delay is 40 msec* (normalized delay of 10), then (from Figure 2) the lowest q which can be used are .92, .84, and .78 for D^{s1}, D^{s2}, and D^h, respectively. When

*Larry Roberts suggested in the PR meeting of December 12-13, 1972, that a desirable average delay for voice, end to end, is 100 msec. Thus we assume that each packet will "travel" independently in the ARPA type network (see PR Temporary Note #5) for an estimated 20 msec, and 40 msec are left for each end of the ALOHA type network.

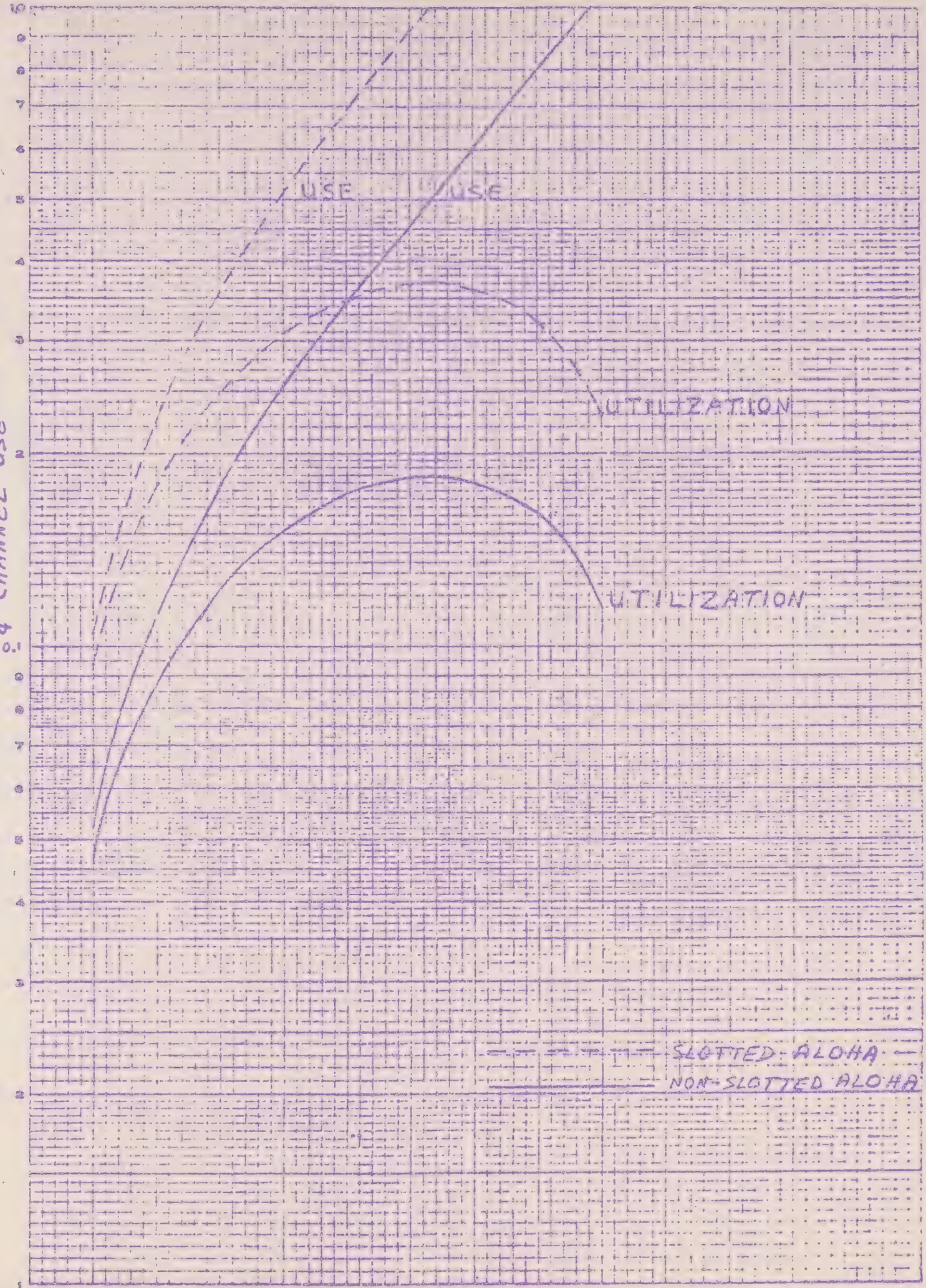
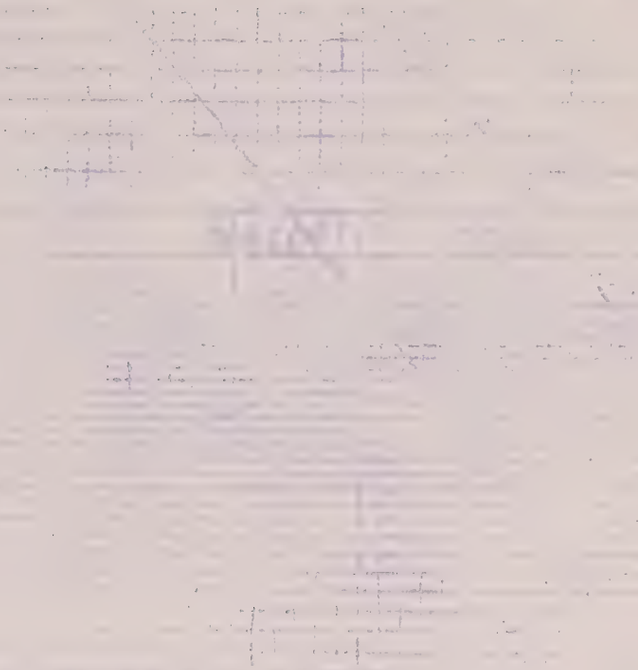

CHANNEL UTILIZATION & CHANNEL USE
USE
USE
UTILIZATION
UTILIZATION
SLOTTED ALOHA
NON-SLOTTED ALOHA
1.0
0.8
0.6
0.4
0.2
0.
P[success] = q

FIGURE 1



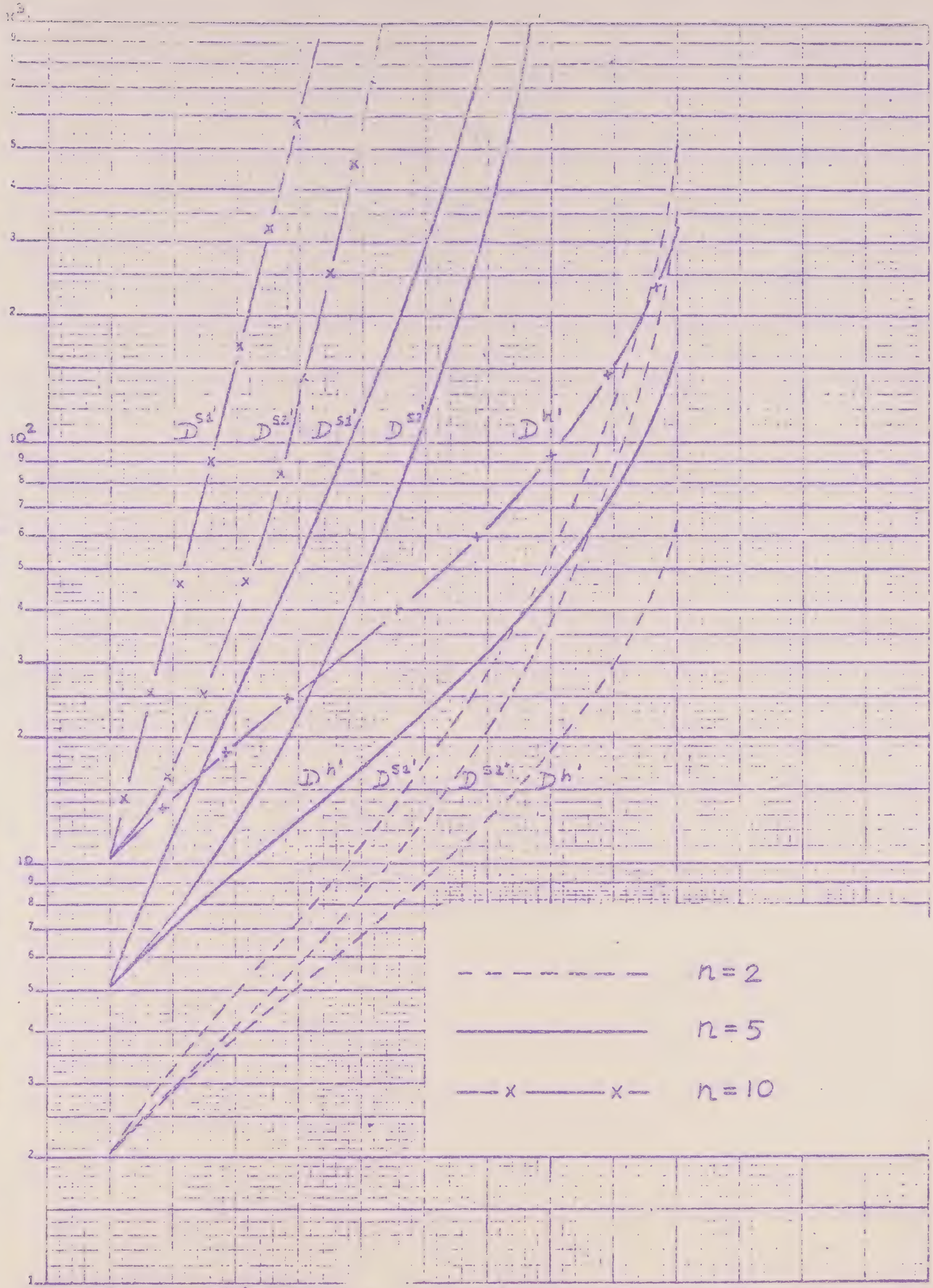


FIGURE 2

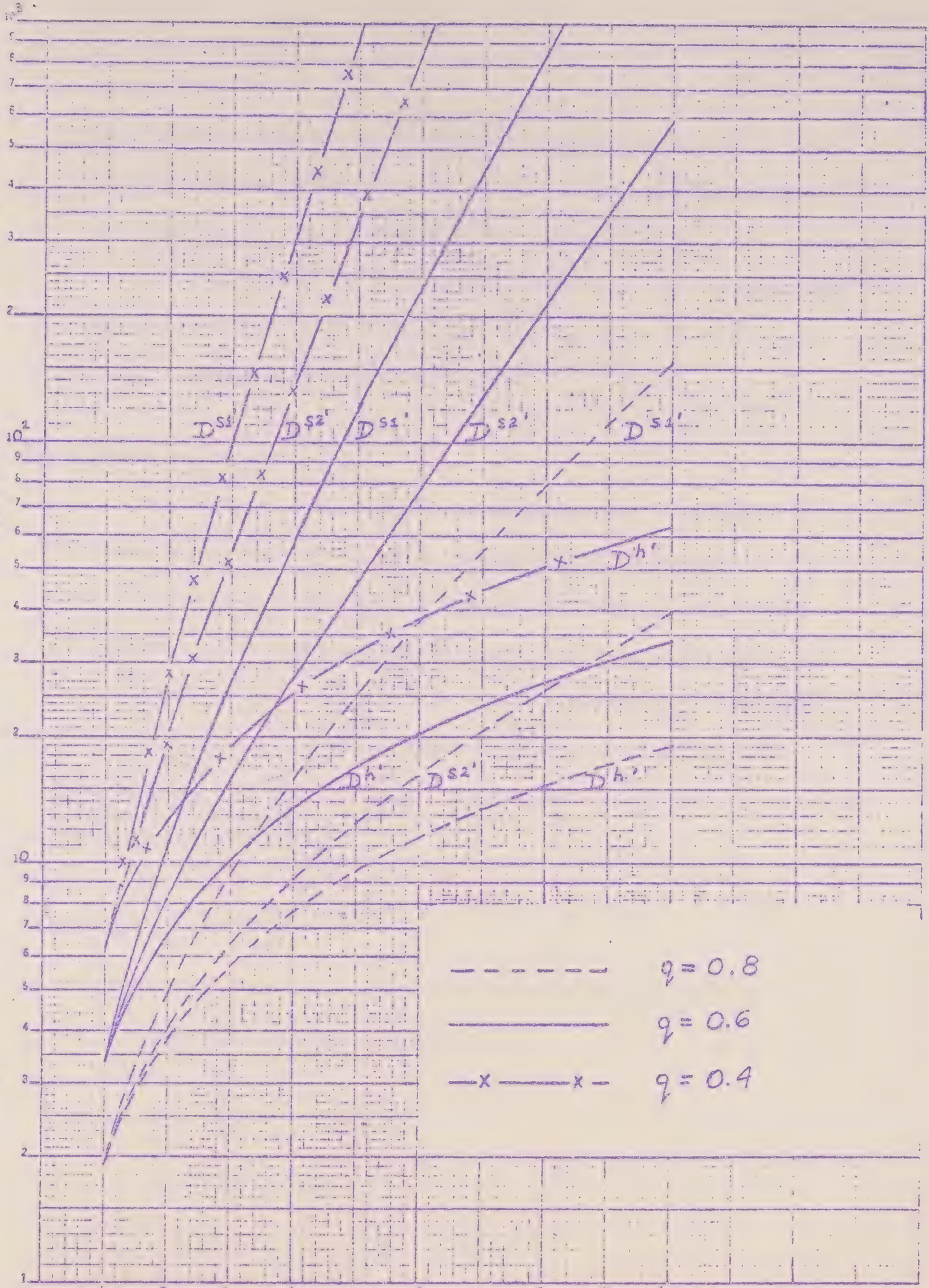


FIGURE 3

a non-slotted random access ALOHA transmission system is used then the maximum effective utilization which can be obtain (see Fig. 1) are 4%, 7.1% and 9.5%, respectively.

Figure 3 shows the normalized delay as a function of the number of hops n, with q as a parameter. Note that $D^{h'}$ is a linear function of n.

IV. q Function of n

In practice q will differ along the path. It seems reasonable to assume that the probability of success, q, will decrease when the packet approaches the station. When a random access ALOHA system is used, the practical range for q is from 1/e for which the effective utilization is maximum to 0.9 for which the utilization is 4.7% and 9.4% for the non-slotted and slotted case, respectively. We take a function of the form:

$$q_i = 0.9 - 0.5 \frac{i}{n} ; \quad i = 1, 2, \ldots n \qquad (9)$$

The normalized average delay as a function of n, with q_i as in Equation (9) is shown in Figure 4.

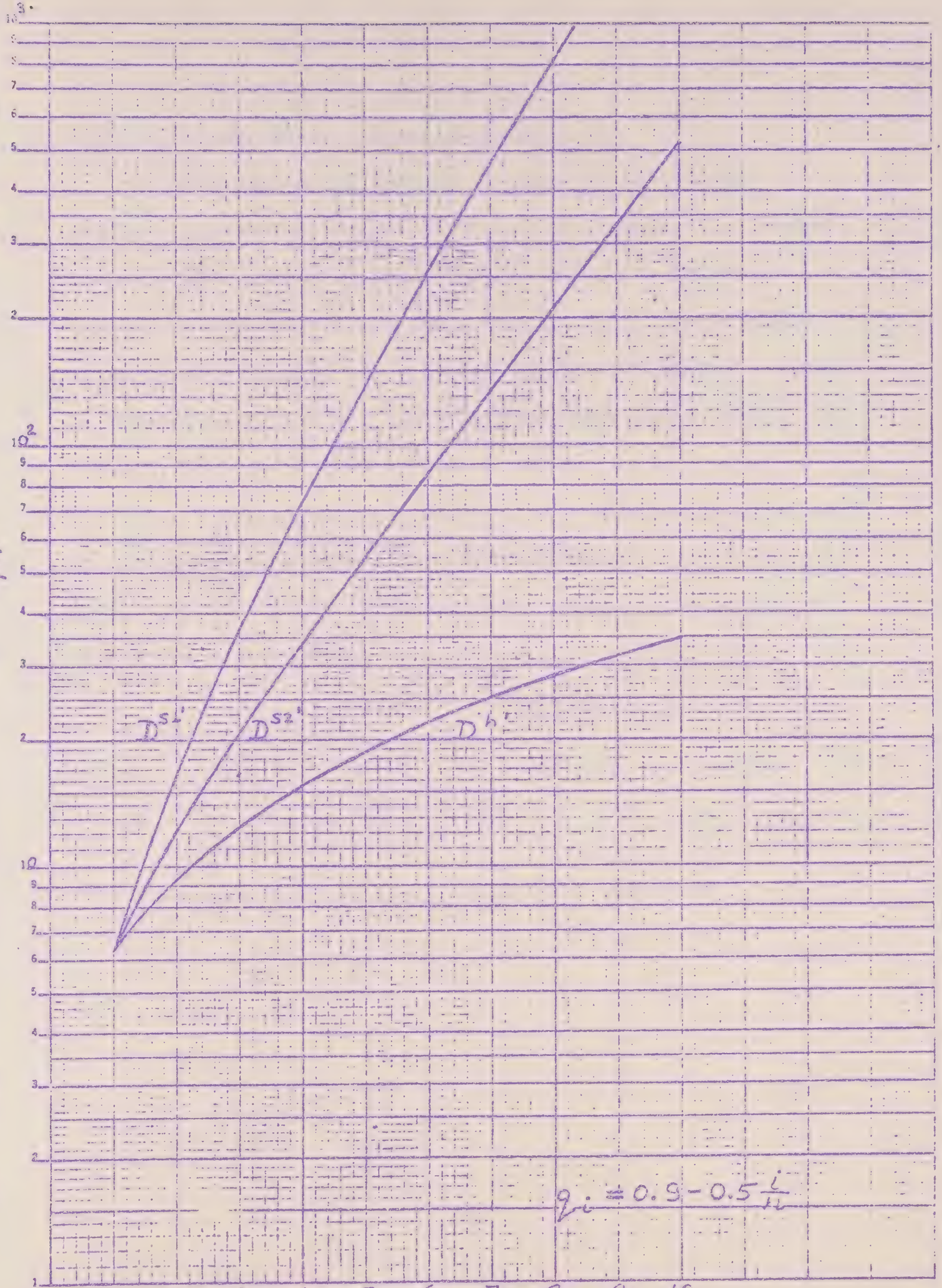
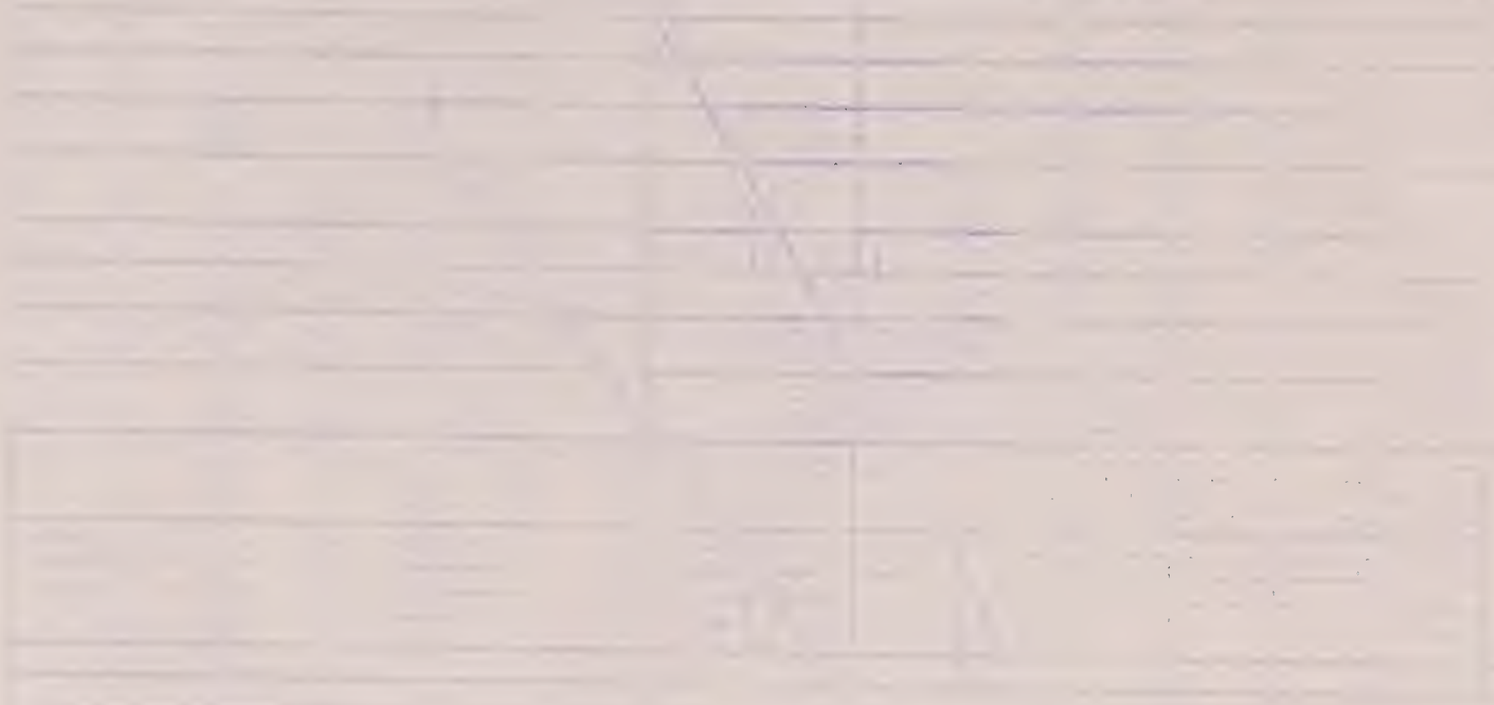


FIGURE 4



V. Utilization for Random Access ALOHA System

In this section we examine the penalty in channel utilization when using end-to-end acknowledgment. We further simplify the problem by using the following assumptions: q has the same value along the path, $\delta = 0$, a slotted random access ALOHA transmission system, that arrivals to each repeater is Poisson with mean δ, arrivals with retransmissions is also Poisson with mean γ, and $T_f = 1$.

Thus, suppose that one wants to use end-to-end acknowledgment in an n-hop system such that the average delay equals that when using hop-by-hop acknowledgment. Equating Equations (6) and (7), and (6) and (8), respectively, one obtains:

$$q^{s1} = (q^h)^{1/n} \tag{10a}$$

$$q^{s2} = \left(\frac{q^h}{n-(n-1)q^h}\right)^{1/n} \tag{10b}$$

When a slotted ALOHA is used, $q=\exp(-\gamma)$, and then:

$$\gamma^{s1} = \frac{\gamma^h}{n} \tag{11a}$$

$$\gamma^{s2} = \frac{1}{n}\, y^{-1} \tag{11b}$$

where

where

$$y = \frac{e^{-\gamma h}}{n-(n-1)e^{-\gamma h}}$$

and consequently:

$$\frac{\rho_{s1}}{\rho_h} = \frac{1}{n} \; e^{\gamma h}\left(1-\frac{1}{n}\right) \qquad (12a)$$

$$\frac{\rho_{s2}}{\rho_h} = \frac{1 \; e^{\gamma h}}{n \cdot \gamma h} \cdot \; y^{1/n} \ln y^{-1} \qquad (12b)$$

Note from (11a) that by using hop-by-hop acknowledgment one can offer n times the rate of packets with retransmissions as in the system with end-to-end acknowledgment and have the same average delay.

The ratios of utilization (Equations (12a) and (12b)) as a function of n are shown in Figure 5, for the case $\gamma h = 0.5$ which is equivalent to 30% utilization in the slotted ALOHA random access system.

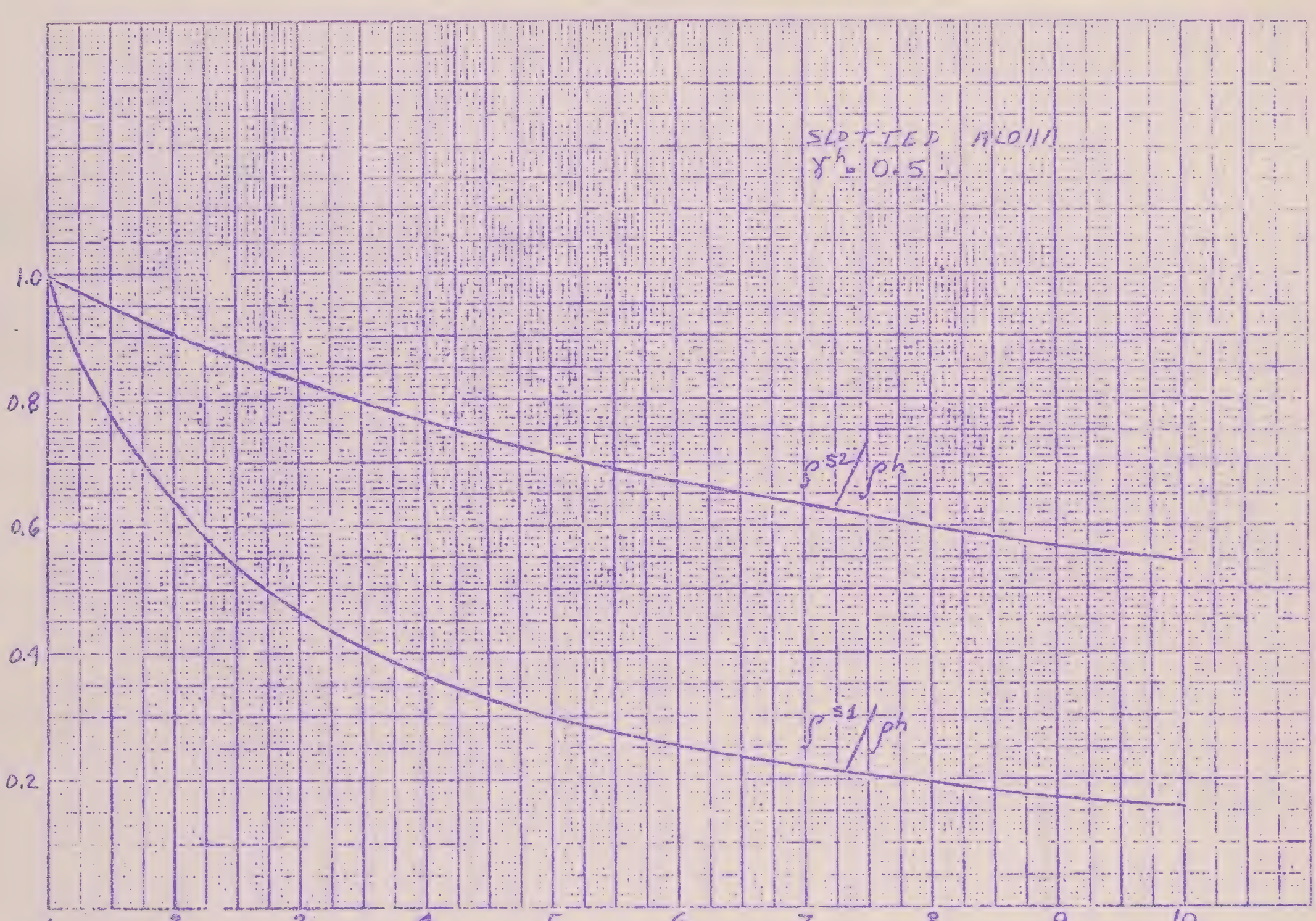
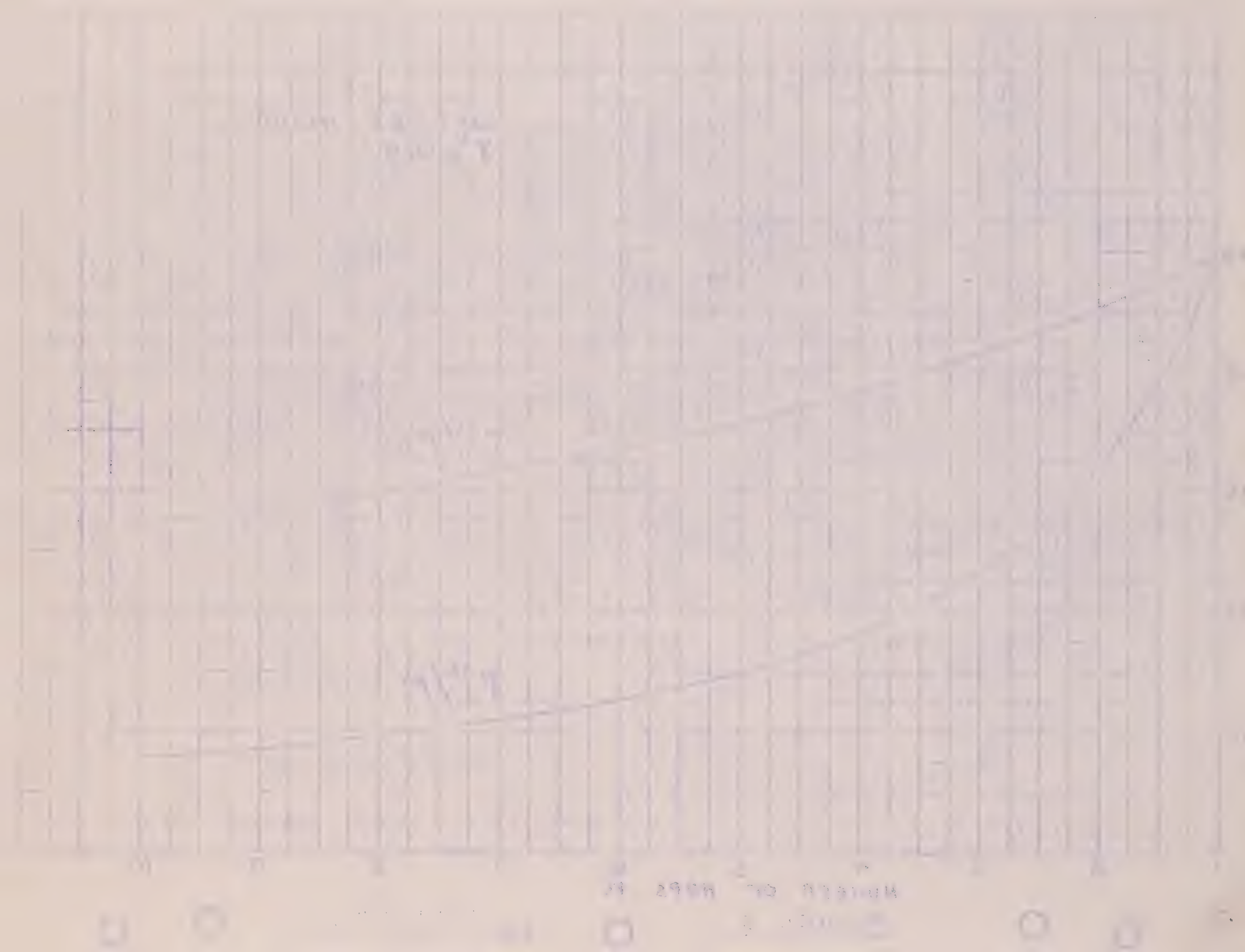

SLOTTED ALOHA
$\gamma^h = 0.5$
ρ^{S2}/ρ^h
ρ^{S1}/ρ^h
1.0
0.8
0.6
0.4
0.2
1
2
3
4
5
6
7
8
9
10
NUMBER OF HOPS n

FIGURE 5



VI. Remarks

(1) In writing D^{s1} it was assumed that the terminal knows its distance in "hops" to the station, thus it knows the minimum waiting time for the acknowledgment to arrive. In practice, however, this may not be true if we assume an operation in which a person types in the packet once and it is automatically retransmitted until a positive acknowledgment has been received. Thus, if this waiting time is set to be a high value then unnecessary delays will result. On the other hand, if it is set to be a small value then in transmitting from a remote area it will generate many copies of the same packet. This, in fact, is the case when operating as described for D^{s2}. Thus, it is the case in which the terminal is adjusted to operate in an urban area and operates in a rural area. Note that this problem is not encountered when using hop-by-hop acknowledgment since the retransmission waiting is always set for one hop.

(2) Another problem, not related to delay, is that if one introduces a maximum number of automatic retransmissions by the terminal, then this number will usually depend, among others, on the number of hops.

VI. Remarks

(1) In writing D_{s1} it was assumed that the terminal knows its distance in "hops" to the station, thus it knows the minimum waiting time for the acknowledgment to arrive. In practice, however, this may not be true if we assume an operation in which a person types in the packet once and it is automatically retransmitted until a positive acknowledgment has been received. Thus, if this waiting time is set to be a high value then unnecessary delays will result. On the other hand, if it is set to be a small value then in transmitting from a remote area it will generate many copies of the same packet. This, in fact, is the case when operating as described for D_{s2}. Thus, it is the case in which the terminal is adjusted to operate in an urban area and operates in a rural area. Note that this problem is not encountered when using hop-by-hop acknowledgment since the retransmission waiting is always set for one hop.

(2) Another problem, not related to delay, is that if one introduces a maximum number of automatic retransmissions by the terminal, then this number will usually depend, among others, on the number of hops.

(3) The time delay in the equations is that for the packets to arrive at the station. It is not the delay for the acknowledgment to be received by the terminal. For the end-to-end acknowledgment scheme a similar (depending on the q_i's in the reverse direction) waiting time is necessary before the acknowledgment is received. The same holds true for hop-by-hop transmission if the acknowledgment from the first repeater is used to stop the retransmission and when the terminal is awaiting a final delay acknowledgment from the station.

(4) In the model considered we set the waiting time for retransmission to be the minimum time for the acknowledgment packet to arrive. In practice, R can be set to be the expected time for the acknowledgment to arrive. If this is done then the model will be further complicated by the introduction of q_i in the backward direction. However, it seems to us that the comparison will lead to similar conclusions.

(3) The time delay in the equations is that for the packets to arrive at the station. It is not the delay for the acknowledgment to be received by the terminal. For the end-to-end acknowledgment scheme a similar (depending on the d_j's in the reverse direction) waiting time is necessary before the acknowledgment is received. The same holds true for hop-by-hop transmission if the acknowledgment from the first repeater is used to stop the retransmission and when the terminal is awaiting a final delay acknowledgment from the station.

(4) In the model considered we set the waiting time for retransmission to be the minimum time for the acknowledgment packet to arrive. In practice, R can be set to be the expected time for the acknowledgment to arrive. If this is done then the model will be further complicated by the introduction of d_j in the backward direction. However, it seems to us that the comparison will lead to similar conclusions.